Segmentation Based Combined Wavelet-Curvelet Approach for Image Denoising

Preety D. Swami [#1], Alok Jain[#2]

[#] Department of Electronics and Instrumentation
Samrat Ashok Technological Institute, Vidisha
RGPV University, Bhopal, India

[1]preetydswami@yahoo.com
[2]alokjain6@rediffmail.com

***Abstract*-This work proposes the use of both the wavelet transform and the curvelet transform for denoising of images corrupted by AWGN. The wavelet reconstruction contains artifacts along the edges in an image. These edges can be captured efficiently by curvelets but curvelets are challenged by smooth regions where artifacts are largely visible. The desirable algorithm for denoising a noisy image should preserve the fine structures in the image and at the same time should not introduce artifacts. In this work the areas containing edges are denoised using curvelet transform while the remaining homogenous regions are recovered through wavelet transform. The areas containing edges and those that do not contain edges are segmented in the space domain by calculating a variance image and then thresholding it. The wavelet and curvelet denoising are inspired by methods in which the wavelet and curvelet coefficients are analyzed statistically and separated into two classes depending upon the probability whether a given coefficient contains a significant noise-free component or not. The algorithm is tested on some images for several noise levels and the results mostly show better performance than the recent methods which employ PDF based models. Also this method provides visually pleasant images as compared to the individual performances of either wavelets or curvelets.**

***Keywords*-Curvelets; Image Denoising; Parameter Estimation; Statistical Modelling; Wavelets**

I. INTRODUCTION

Degradation of an image during acquisition or transmission is due to the introduction of artifacts like additions of noise, blurring, distortion etc. Image denoising, a significant topic of image restoration, intends to remove the noise while retrieving the defining features in the image. Removal of noise can be done by linear processing such as Wiener filtering. Non-linear techniques employ thresholding or shrinking of transformed coefficients. Many multiscale transforms have been developed in recent years which have made efforts in improving the results of denoising. Application of wavelets is very popular in this field and an ample amount of work has been contributed by various authors [1-6]. Several other transforms with directional multiresolution bases have also been introduced such as the dual-tree complex wavelet transform (DT-CWT) [7], wedgelet [8], bandlet [9], ridgelet [10], shearlet [11], contourlet [12], platelet [13], and curvelet [14]. Some authors [15, 16] have designed adaptive threshold functions in which thresholding is done using neural network approach. Considerable researches have been done in the field of sparse and redundant representation modelling [17, 18] and are potentially very powerful. In these approaches data characteristics are learnt either from the image itself or from some existing database and are then used as dictionaries.

Image corrupted by additive noise can be restored efficiently if the statistical difference between image signal and noise signal can be computed. An image is a combination of high-contrast region like edges and low-contrast areas which have a smooth appearance. Wavelets have dominated this field for the past few years but they failed to smoothly restore the edges, which is necessary for better perception of an image. The curvelet transform [14] finds an advantage over wavelets in that, it requires fewer coefficients to represent edges and gives a better perception. But at the same time it suffers from occurrence of visual artifacts in the low-contrast region. Thus for two different areas (edges/smooth) in an image two different denoising methods can be employed so that combined attributes of both the transforms benefit denoising.

In most of the image processing problems the marginal statistics of image are generally modelled as generalized Gaussian densities. Simple denoising algorithms based on shrinking of transform coefficients [14, 19] are being replaced by methods in which relationships between coefficients are also taken into account. Some algorithms consider the coefficients to be independent, while others take advantage of correlation of intra-band, inter-scale and inter-band/inter-orientation transform coefficients [20-23].

The approach for denoising in this work is influenced by the method [24] of dealing with speckle noise removal. The method denoises an image by simply thresholding the transform coefficients without taking into account the correlation of transform coefficients. In this work the image is segmented into regions containing edges and smooth areas. Both wavelet and curvelet transforms of the noisy image are computed and the coefficients are processed by using local spatial context based modelling. Thereafter, the edges in the image are substituted by pixels of curvelet denoised image while the remaining smooth regions are substituted by the wavelet denoised pixels. A comparison with some of the best wavelet-based and curvelet-based denoising methods illustrates the competitiveness of this approach.

The paper is organized as follows. Section 2 is devoted to the study of spatial adaptive threshold selection using context modelling. The developed image denoising algorithm is introduced in Section 3. In Section 4 the implementation details and the simulation results are presented. Finally, the concluding remarks are stated in Section 5.

II. CONTEXT MODELLING FOR SPATIAL ADAPTIVITY

Noise reduction is the result of shrinking of noisy coefficients in the transform domain. The transformed coefficients which contain a significant noise free component should be reduced to a lesser amount. On the other hand, coefficients that are dominated by noise should be suppressed to zero in an ideal case. Analysis of image statistics shows that images do not reflect Gaussian nature. Hence, for the purpose of denoising, spatially adaptive models are used by assuming locally Gaussian nature of images. Arrival of multiscale and multioriented transforms has facilitated the use of local adaptation of thresholds according to the spatially changing statistics of images. Chang et al. [25] proposed a spatially adaptive wavelet threshold method which is based on context modelling. In context modelling, grouping of those pixels is done which are identical in nature, irrespective of their spatial adjacency. Thereafter, the model is used to estimate the parameters for each wavelet coefficient. Here each wavelet coefficient is considered to be a random variable of Generalized Gaussian Distribution with parameters *s* (scale parameter) and *p* (shape parameter).

Assume that the original image is $\{W[i,j],\ i,j=1,2,....N\}$, and is contaminated with additive white Gaussian noise of zero mean and variance σ^2. Thus noisy image is

$$Y[i,j] = W[i,j] + \varepsilon[i,j], \qquad i,j = 1,2,.....N \tag{1}$$

Where $N \times N$ is the dimension of the image. Let *y* and *w* denote the matrix of wavelet coefficients of *Y* and *W* respectively. Estimation of σ_w, of each wavelet coefficient $y[i,j]$ in a particular subband is done by first considering its nearest neighbourhood in the same subband along with its parent coefficients. The context then is the weighted average of the absolute values of its neighbours

$$z[i,j] = \varpi^t u_{ij} \tag{2}$$

u_{ij} is a vector containing absolute values of those elements which form a neighbourhood of $y[i,j]$. Using the least squares estimate, the weight can be calculated as

$$\varpi_{LS} = \arg \min_{\varpi} \sum_{i,j} \left(|y[i,j]| - \varpi^t u_{ij} \right)^2 \tag{3}$$

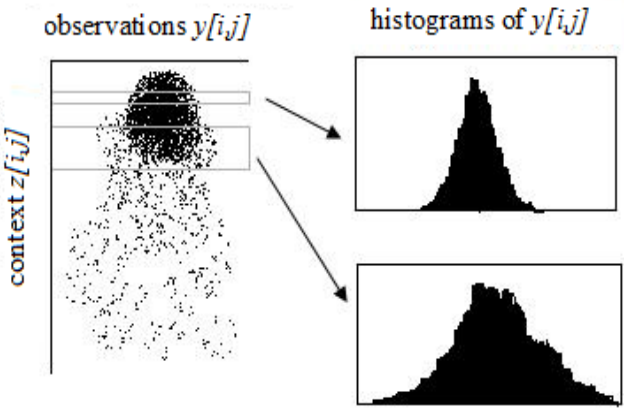

Fig. 1 Plot of context $z[i,j]$ vs. noisy wavelet coefficients $y[i,j]$. Both the windows contain the same number of points leading to the conclusion that local variability can be sufficiently estimated by the measure of context.

The coefficients whose context variables are in proximity with $z[i,j]$ will determine the variance of the random variable $y[i,j]$. This can be depicted from Fig. 1 which shows the relation of $y[i,j]$ with $z[i,j]$. Smaller spread in $y[i,j]$ is observed for a context $z[i,j]$ of small interval. Correspondingly, when $z[i,j]$ is large it results in larger spread of $y[i,j]$.

Variance of any given coefficient $y[i_0, j_0]$ can be estimated by placing a window around $z[i_0, j_0]$. Those points $y[i,j]$ whose context, $z[i,j]$ lie inside the window will then estimate the variance. This is done by taking L closest points above $z[i,j]$ and, L closest points below. The estimate of variance σ_w^2 is then

$$\hat{\sigma}_w^2[i_0, j_0] = \max\left(\frac{1}{2L+1} \sum_{[k,l] \in B_{i_0 j_0}} y[k,l]^2 - \sigma^2, 0 \right) \tag{4}$$

The points $\{y[i,j]\}$, whose context lie inside the window form the set $B(i_0, j_0)$. Since the noise is independent of the signal, hence the noise variation σ^2 is subtracted from the noisy observation $\{y[i,j]\}$.

The threshold at location $[i_0, j_0]$ is given by

$$T_B[i_0, j_0] = \frac{\sigma^2}{\hat{\sigma}_w[i_0, j_0]} \tag{5}$$

For every location $[i,j]$, the spatially adaptive threshold $T_B[i,j]$ can thus be calculated. Noise variance σ^2, can be estimated by median estimator in the highest subband of wavelet transform

$$\hat{\sigma} = Median\ (\ |y[i,j]|\ ) / 0.6745 \tag{6}$$

III. PROPOSED APPROACH

The aim of denoising is to optimize the mean square error between the original image *W* and $\hat{W}$, which is the image estimated through denoising. The approach of denoising in this work is firstly to divide the image into two regions, namely the flat areas and the edges, using any segmentation scheme. The flat areas and the edges are denoised using spatially adaptive Bayesian estimators in wavelet and curvelet domain respectively. Using a spatial adaptive Bayesian estimator is not only faster but it is simpler to implement. Also it outperforms the recent complex procedures which are based on HMT and MRF models. Finally the two images are fused together. The proposed algorithm thus is a four step process, (1) segmentation (2) wavelet denoising (3) curvelet denoising (4) fusion.

Step 1: Image Segmentation

Edge points in the noisy image are differentiated from smooth regions by segmenting the noisy image in the space domain. Segmentation of the image into two classes can be done by calculating a variance map of the image and then thresholding it by employing a suitable threshold [26]. The variance map of an image is the collection of variances of each pixel of the image. This map is calculated by taking an

odd-sized square window around a centre pixel of intensity x_i. The window is then made to slide around every pixel to give the complete variance map. Within the window the variance is calculated as

$$\sum_{i=1}^{n_p}\left(x_i-\mu\right)^2\Big/(n_p\text{-}1) \tag{7}$$

Where n_p, is the number of pixels inside the window, and μ is the mean map of the image in which each pixel is the mean of the pixel values inside the window, centred around the pixel whose variance is to be calculated. Fig. 2 shows the result of segmentation of some standard images used in this work.

Step 2: Wavelet Denoising

The wavelet denoising method used in this work is the *ProbShrink* [22]. In *ProbShrink,* the amount of shrinking of the noisy coefficients in the same subband depends on their spatial position and their local surrounding. The noisy input image corrupted with additive white Gaussian noise is transformed using the wavelet transform. For *n* coefficients in each subband of the wavelet transform we have

$$y_i = w_i + \varepsilon_i, \qquad i = 1.....n \tag{8}$$

Where, y_i are noisy wavelet coefficients, w_i are noise-free wavelet coefficients and ε_i are independent normal random variables of zero mean and variance σ^2.

For the noise-free subband data the prior assumed is the generalized Laplacian [25] which can be expressed as

$$f(w) \propto \mathrm{e}^{-s|w|^p} \tag{9}$$

Where s and p are the scale and shape parameters respectively and can be estimated from the noisy histogram. Given a generalized Laplacian signal corrupted with AWGN of strength σ, its second moment (variance) and the fourth moment are

$$\begin{aligned}\sigma_y^2 &= \sigma^2 + \left(\Gamma\left(3/p\right)/s^2\Gamma\left(1/p\right)\right),\\ m_{4,y} &= 3\sigma^4 + \left(6\sigma^2\,\Gamma\left(3/p\right)/s^2\Gamma\left(1/p\right)\right) + \Gamma\left(5/p\right)/s^4\Gamma\left(1/p\right)\end{aligned} \tag{10}$$

Where $\Gamma(x)$, is the gamma function. For each subband, σ_y^2 and $m_{4,y}$ can be calculated using statistical measures. Using (10), the kurtosis k_y, of generalized Laplacian distribution can be calculated as

$$k_y = \left(m_{4,y} + 3\sigma^4 - 6\sigma^2\sigma_y^2\right)\Big/\left(\sigma_y^2 - \sigma^2\right)^2 \tag{11}$$

Knowing the kurtosis k_y, the parameter p can be solved numerically from

$$k_y = \Gamma\left(5/p\right)\Gamma\left(1/p\right)\Big/\left(\Gamma\left(3/p\right)\right)^2 \tag{12}$$

The parameter s can now be calculated as

$$s = \left(\left(\sigma_y^2 - \sigma^2\right)\left(\Gamma\left(1/p\right)/\Gamma\left(3/p\right)\right)\right)^{-1/2} \tag{13}$$

These parameters are then employed to shrink the wavelet coefficients. Inside a subband, shrinking of noisy coefficients which are equal in magnitude is not the same but depends upon the coefficients' spatial location and local surrounding. The shrinkage rule for each spatial position l of a coefficient is

$$\hat{w}_l = \frac{\psi\eta_l\xi_l}{1+\psi\eta_l\xi_l}\,y_l \tag{14}$$

Where ψ is the probability of usefulness of each coefficient with respect to the global statistical properties of the coefficient and is given by

$$P(H_1)/P(H_0) = \left(1-\Gamma_{inc}\left((sT)^p, 1/p\right)\right)\Big/\Gamma_{inc}\left((sT)^p, 1/p\right) \tag{15}$$

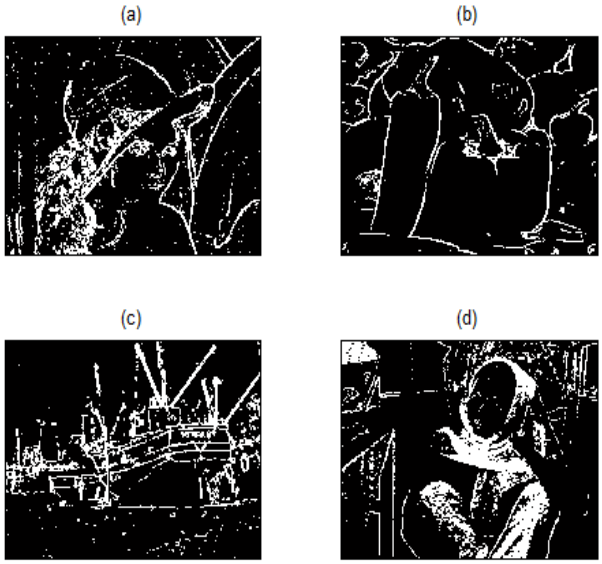

Fig. 2 Segmentation results for some standard images used in this work. (a) Lena (b) Peppers (c) Boats (d) Barbara

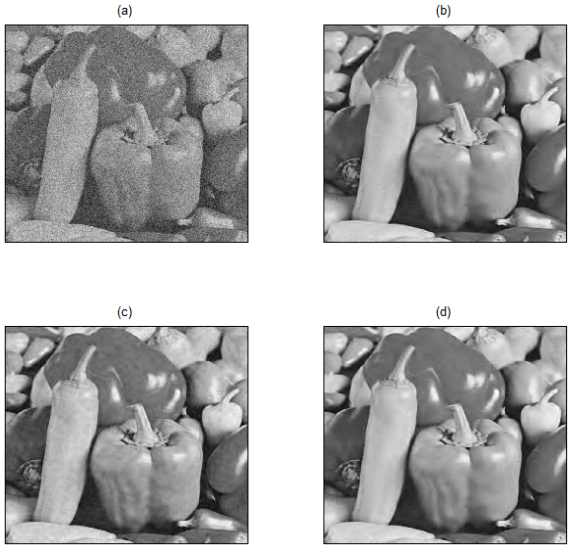

Fig. 3 (a) Noisy image with PSNR = 18.58 dB (σ = 30). (b) Wavelet denoised image using ProbShrink method, PSNR = 28.35 dB. (c) Curvelet denoised image using ProbShrinkCurve method, PSNR = 29.73 dB. (d) Denoised image using the proposed method, PSNR = 31.10 dB.

Assume H_1 is chosen to denote the presence of "signal of interest" and H_0 to represent "absence of signal of interest". Then for a given coefficient w and a specific threshold T, H_0 : $w \le T$ and H_1 : $w > T$. Thus $P(H_1)$ is the probability that noise free signal is present in a given coefficient which amounts to the area under $f(w)$ exceeding the threshold T, and

$P(H_0)$ is the probability that noise free signal is absent in a given coefficient. $\Gamma_{inc}(x,a)$, is the incomplete gamma function.

η_l in (14) gives the information about the usefulness of the coefficient itself. It is the ratio of conditional densities of noisy coefficients

$f(y_l \mid H_1)/f(y_l \mid H_0)$. These conditional densities are given by

$$\begin{aligned} f(y_l \mid H_0) &= \int_{-\infty}^{\infty} \Phi(y_l - w;\sigma) f(w \mid H_0)\, dw \\ f(y_l \mid H_1) &= \int_{-\infty}^{\infty} \Phi(y_l - w;\sigma) f(w \mid H_1)\, dw \end{aligned} \tag{16}$$

Where $\Phi(y;\sigma)$ is the zero mean Gaussian density and $f(w \mid H_0)$ and $f(w \mid H_1)$ are conditional densities of noise-free coefficients given by

$$\begin{aligned} f(w \mid H_0) &= \begin{cases} W_0 \exp\left(-s|w|^p\right) & if \ |w| \le T \\ 0 & if \ |w| > T \end{cases} \\ f(w \mid H_1) &= \begin{cases} 0 & if \ |w| \le T \\ W_1 \exp\left(-s|w|^p\right) & if \ |w| > T \end{cases} \end{aligned} \tag{17}$$

Where W_0 and W_1 are normalization constants

$$\begin{aligned} W_0 &= sp/2\Gamma(1/p)\Gamma_{inc}\left((sT)^p, 1/p\right) \\ W_1 &= sp/2\Gamma(1/p)\left(1-\Gamma_{inc}\left((sT)^p, 1/p\right)\right) \end{aligned} \tag{18}$$

ξ_l, in (14) is measured from the local surroundings and is given by $f(z_l \mid H_1)/f(z_l \mid H_0)$, in which z_l is the local spatial activity indicator. For each y_l in the same subband, z_l is the average of coefficient magnitudes inside a small window centred at y_l.

Inverse wavelet transformation of the coefficients which are shrunk by the rule in (14) yields the denoised image.

Step 3: Curvelet Denoising

Curvelet basis elements follow a parabolic scaling relation and are supported compactly in the frequency as well as in the spatial domain. The geometric multiscale features [27] of curvelets are better than that of wavelets which are advantageous when edges in an image are to be neatly reconstructed. Large numbers of wavelet coefficients are required to accurately capture edges in an image whereas curvelet transform requires only a few non-zero coefficients for their reconstruction.

The denoising method used in this work is adopted from the *ProbShrinkCurve* method [23] which is inspired by the *ProbShrink* method. Same steps will be followed as were for the wavelet denoising. By analysing the curvelet coefficients statistically, the coefficients containing a noise-free component are separated from those that do not contain a signal of interest.

In order to suppress pseudo-Gibbs artifacts in the curvelet denoised image, post processing of curvelet shrinkage result can be done [28]. For this purpose the total variation (TV) can be used to regularize images without smearing away the object boundaries. This work employs TV minimization method in which ROF (Rudin-Osher-Fatemi) model is used to minimize the functional [29]

$$F(u) = \min_u \quad TV(u) + \frac{\lambda}{2}\|f_c - u\|^2 \tag{19}$$

Where the output *u,* is the piecewise smooth component of the input f_c, which is the curvelet denoised image. λ, is the regularization parameter, and $TV(u)$ is the total variation of *u*. The minimization is solved using Chambolle's method [30].

Step 4: Fusion

Fusion of the two images denoised by wavelet and curvelet denoising is done on the basis of segmentation. In the segmented image, the pixels that correspond to edges are replaced by the corresponding pixels of the curvelet denoised image. Remaining pixels pertaining to smooth regions are substituted by the respective pixels of the wavelet denoised image.

IV. SIMULATION RESULTS

Test results of the proposed algorithm on the images of *Peppers, Lena and Boat* is reported. All the images are of size 512×512. For creating the variance map of the image a square-shaped window of size 5×5 is employed. For segmentation into two classes i.e. edges and smooth areas, the threshold parameter was empirically calculated and set to 2500. Four levels of wavelet decomposition using *symmlet*, with three orientations per level are employed for wavelet denoising the noisy image and the threshold is set to $1.5\times\sigma$. The noisy image is curvelet denoised using the digital curvelet transform based on wrapping of Fourier coefficients [31]. Curvelet decomposition is done at 4 scales and number of orientations at the coarsest curvelet level is selected as 16. At the finest scale, there is a choice between wavelet and curvelet decomposition for which curvelet decomposition is opted in this work. Through experiments the threshold for curvelet denoising is set to $0.5\times\sigma$. The parameter λ for post processing of curvelet denoised image using TV minimization is selected as 0.7.

The comparison of the proposed method with some recent methods, in terms of PSNR, is listed in Table 1. In comparison, the wavelet- based denoising methods selected are the *BiShrink* [21], trained adaptive dictionaries [17], the *ProbShrink* [22], the *SVGSM* [32] and the *MPGSM* [33]. *BiShrink* is a locally adaptive denoising algorithm which uses the bivariate shrinkage function. Elad et al. [17] presented a method in which instead of choosing a predefined set of basis functions, learning is done through a dictionary. The dictionary is trained either on the noisy image or on an image data-base. In the spatial adaptive version of *ProbShrink* method, each wavelet coefficient is shrunk by an amount which depends upon its location and local surrounding in the subband. *SVGSM* model is based on BLS estimator using space variant GSM. In *MPGSM*, the neighborhoods of subbands are modeled by linearly projecting the MGSM model. Results of two curvelet denoising methods, the *DCuT* [14] and *ProbShrinkCurve* [23] are also compared. *DCuT* employs simple hard thresholding of curvelet coefficients, and the *ProbShrinkCurve* is an adaptation of *ProbShrink* method to curvelets by analyzing the statistical behavior of curvelet coefficients. From the tabular results it is observed that the

TABLE I
PSNR VALUES IN DECIBELS FOR SEVERAL DENOISING METHODS

				$PSNR_{out}$						
	σ	$PSNR_{in}$	Bishrink in [21]	DCuT in[14]	Learned Dictionaries in [17]	ProbShrink in [22]	SVGSM in [32]	ProbShrinkCurve in [23]	MPGSM in [33]	Proposed
Peppers	5	34.14	-	36.43	37.78	-	37.72	36.70	37.62	38.35
	10	28.13	-	33.32	34.32	33.90	34.24	34.22	34.77	35.90
	15	24.61	-	31.37	32.37	31.81	32.18	-	33.32	34.14
	20	22.10	-	30.06	30.92	30.30	30.67	31.63	32.25	32.98
	25	20.22	-	29.25	29.84	29.23	29.50	-	31.38	31.94
	30	18.58	-	28.58	-	28.35	-	29.73	-	31.10
	50	14.15	-	26.59	25.90	-	26.11	26.81	-	27.95
Lena	5	34.14	38.01	36.60	38.60	-	38.55	37.86	38.63	38.87
	10	28.13	35.34	33.78	35.47	35.24	35.66	35.20	35.74	35.96
	15	24.61	33.67	32.22	33.90	33.46	33.96	33.38	34.04	34.06
	20	22.10	32.40	31.07	32.66	32.20	32.71	32.02	32.80	32.70
	25	20.22	31.40	30.15	31.69	31.21	31.72	30.97	31.81	31.75
	30	18.58	30.54	29.32	-	30.33	-	30.01	-	30.65
	50	14.15	-	27.09	27.79	-	28.61	27.14	-	27.71
Boat	5	34.14	-	33.59	37.22	-	-	36.19	-	35.35
	10	28.13	33.10	30.57	33.64	33.25	-	33.12	-	33.65
	15	24.61	31.36	29.10	31.73	31.32	-	31.23	-	31.78
	20	22.10	30.08	28.15	30.36	29.93	-	29.93	-	30.47
	25	20.22	29.06	27.43	29.28	28.89	-	28.88	-	29.34
	30	18.58	28.31	26.80	-	28.04		28.01	-	28.66
	50	14.15	-	25.08	25.95	-	-	25.35	-	25.95

TABLE II
COMPARISON WITH RECENT METHODS EMPLOYING NEURAL NETWORKS: SNR AND PSNR VALUES ARE IN DECIBELS

			$PSNR_{out}$				SNR_{out}		
	σ	$PSNR_{in}$	WT-TNN in [15]	PSO with bior6.8 in [16]	Proposed	SNR_{in}	WT-TNN in [15]	PSO with bior6.8 in [16]	Proposed
Lena	10	28.13	34.24	34.34	35.96	13.58	19.70	19.80	21.42
	20	22.10	30.70	30.94	32.70	7.59	16.26	16.40	18.11
Barbara	10	28.13	31.69	31.81	33.72	13.42	18.32	18.45	18.67
	20	22.10	27.57	27.64	30.12	7.40	14.20	14.25	14.74

proposed combined wavelet-curvelet method outperforms other methods in most cases. The qualitative difference between the methods which use only the wavelets or the curvelets and the proposed approach is represented in Fig. 3.

In this figure, blurring of edges is observed when denoising is done by using wavelets. The curvelet denoised image reproduces the edges properly but the image is not visually pleasant due to the presence of curvelet-like artifacts in the smooth regions. The proposed method can denoise the image with clean edges and artifact-free homogenous areas.

In Table 2, comparison of proposed approach is done by two recent neural-network-based thresholding methods: WT-TNN [15] and PSO-based technique [16]. WT-TNN (wavelet transform-based thresholding neural network) methodology is independent of the distribution model of the wavelet coefficients. It learns the parameters of thresholding function and threshold value simultaneously in each sub-band of wavelet transform. It has been proven to outperform several thresholding techniques like the soft, hard and garrote thresholding. The limitation of WT-TNN is the requirement that, for fast convergence of learning process, proper values of thresholding parameters need to be initialized. An improvement of this approach is suggested in [16] which utilizes PSO (particle swarm optimization) algorithm for learning the thresholding parameters. The methods in [16] and

[17] are compared with two types of filters, the 'db8' wavelet filter and the 'bior6.8' wavelet filter. Results of denoising when employing 'bior6.8' are better as compared with that of 'db8' filter. Proposed algorithm is compared with the best available results of filtering with 'bior6.8' filter. For standard images of Lena and Barbara the proposed algorithm proves to yield significantly superior results in terms of SNR and PSNR.

V. CONCLUSIONS

This work proposes a combination of wavelets and curvelets for denoising of images. Curvelet transform is used for accurate preservation of edges, while wavelet transform is used for obtaining artifact free smooth areas. The differentiation between edges and smooth areas is done via segmentation in the received noisy image. The method of denoising for wavelets and curvelets is similar in that the local statistics and intrascale dependencies of wavelet and curvelet coefficients are analyzed. The algorithm is compared with many other recent methods and significantly better results are obtained as compared to the using of either the wavelets only or the curvelets.

REFERENCES

[1] D. L. Donoho and I. M. Johnston, "Ideal spatial adaptation via wavelet shrinkage," Biometrika, 81(3), pp. 425- 455, 1994.

[2] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no.7, pp. 674-693, 1989.

[3] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in Proc. ICIP, 1996, pp. 379-382.

[4] M. S. Crouse, R. D. Nowak and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," IEEE Trans. Signal Processing, vol. 46, pp. 886-902, 1998.

[5] F. Abramovich, T. Sapatinas and B. W. Silverman, "Wavelet thresholding via a Bayesian approach," J. R. Statist. Soc., vol. 60, pp. 725-749, 1998.

[6] H. Chipman, E. Kolaczyk and R. McCulloch, "Adaptive Bayesian wavelet shrinkage," J. Amer. Statist. Assoc., 92(440), pp. 1413-1421 1997.

[7] N. G. Kingsbury, "Image processing with complex wavelets," Philosophical Tranactions of Royal Society London A, vol. 357, no.1760, pp. 2543-2560, 1999.

[8] D. Donoho, "Wedgelets: nearly minimax estimation of edges," Ann. Statistics, vol. 27, no. 3, pp. 859-897, 1999.

[9] E. Le Pennec and S. Mallat, "Sparse geometrical image approximation with bandlets," IEEE Trans. Image Processing, vol. 14, no. 4, pp. 423-438, 2005.

[10] G. Y. Chen and B. Kegl, "Image denoising with complex ridgelets," Pattern Recognition, vol. 40, pp. 78-585, 2007.

[11] D. Labate, W. Q. Lim, G. Kutyniok and G. Wiess, "Sparse multidimensional representation using shearlets," in SPIE Conference Proceedings, 2005, vol. 5914, pp. 254-262.

[12] R. Eslami and H. Radha, "Translation-invariant contourlet transform and its application to image denoising," IEEE Trans. Image Processing, vol. 15, no.11, pp. 3362-3374, 2006.

[13] R. Willet and K. Nowak, "Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging," IEEE Trans. Med. Imaging, vol. 22, no. 3, pp. 332-350, 2003.

[14] J. L. Starck, E. J. Candes and D. L. Donoho, "The curvelet transform for image denoising," IEEE Trans. Image Processing, vol. 11, no. 6, pp. 670-684, 2002.

[15] M. Nasri and H. N. Pour, "Image denoising in the wavelet domain using a new adaptive thresholding function," Elsevier J. Neurocomputing, vol. 72, pp. 1012-1025, 2009.

[16] G. G. Bhutada, R. S. Anand and S. C. Saxena, "PSO-based learning of sub-band adaptive thresholding function for image denoising," Springer Journal of Signal Image and Video Processing, DOI 10.1007/s11760-010-0167-7, 2010.

[17] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Processing, vol. 15, no. 12, pp. 3736-3745, 2006.

[18] M. Elad, M. A. T. Figueiredo and Y. Ma, "On the role of sparse and redundant representation in image processing," IEEE Proceedings - Special Issue on Applications of Sparse Representations and Compressive Sensing, vol. 98 no. 6, pp. 972-982, 2010.

[19] D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inform. Theory, vol. 41, no. 3, pp. 613-627, 1995.

[20] Pizurica, W. Philips, I. Lemahieu and M. Acheroy, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," IEEE Trans. Image Processing, vol. 11, no. 5, pp. 545-557, 2002.

[21] L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," IEEE Signal Processing Lett., vol. 9, no. 12, pp. 438-441, 2002.

[22] Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," IEEE Trans. Image Process., vol. 15, no. 3, pp. 654-665, 2006.

[23] L. Tessens, A. Pizurica, A. Alecu, A. Munteanu and W. Philip, "Context adaptive image denoising through modeling of curvelet domain statistics," Journal of Electronic Imaging, vol. 17, no. 3, 033021-1—033021-17, 2008.

[24] B. Saevarsson, J. R. Sveinsson and J. A. Benediktsson, "Speckle reduction of SAR images using adaptive curvelet domain," in IGARSS Conference Proceedings, 2003, pp. 4083-4085.

[25] S. G. Chang, B. Yu and M. Vetterli, "Spatially adaptive wavelet thresholding with context modelling for image denoising," IEEE Trans. Image Process., vol. 9, no. 9, pp. 1522-1531, 2000.

[26] S. G. Chang and M. Vetterli, "Spatial adaptive wavelet thresholding for image denoising," in ICIP Conference Proceedings,1997, pp. 374-377.

[27] E. J. Candes and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with C2 singularities," Comm. Pure and Appl. Math., vol. 57, no. 2, pp. 219-266, 2004.

[28] J. Ma and G. Plonka, "The curvelet transform: A review of recent applications," IEEE Sig. Process. Mag. vol. 27, no. 2, pp. 118-133, 2010.

[29] S. Durand, J. Froment, "Reconstruction of wavelet coefficients using total variation minimization," SIAM J. Sci. Comput., vol. 24, pp. 1754-1767, 2003.

[30] Chambolle, "An algorithm for total variation minimization and applications," Journal of Mathematical Imaging and Vision, vol. 20, pp. 89-97, 2004.

[31] E. Candes, L. Demanet, D. Donoho and L. Ying, "Fast discrete curvelet transforms," SIAM Multiscale Model. Simul., vol. 5, no.3, pp. 861-899, 2006.

[32] J. A. Guerrero-Colon, L. Mancera and J. Portilla, "Image restoration using space-variant Gaussian scale mixtures in overcomplete pyramids," IEEE Trans. Image Process., vol. 17, no. 1, pp. 27-41, 2008.

[33] Goossens, A. Pizurica and W. Philips, "Image denoising using mixtures of projected Gaussian scale mixtures," IEEE Trans. Image Process., vol. 18, no. 8, pp. 1689-1702, 2008.